RESEARCH PAPER OPEN ACCESS

# Areca nut Detection and Classification using Segmentation and Computer Vision

Satheesha K M, Mahesh Tubaki
*Electronics and Communication Engg, KPT - Mangalore*
*Electronics and Communication Engg, Govt Polytechnic - Belagavi*

**ABSTRACT**

There are various computer-based technologies for other crops, but no computer vision-based sophisticated technology for recognizing an areca nut's grade, variety, or illnesses. In this study, a unique approach for categorizing areca nuts into two classes based on color is suggested. There are three phases to the suggested method: (i) Segmentation; (ii) Masking; (iii) Classification, a color space conversion is performed on the RGB picture. For successful areca nut segmentation, three sigma control limits are applied to the picture. The color space of the areca nut is simulated using three sigma control limits, which covers the most important fluctuation of the areca nut's color components. The color components' higher and lower boundaries are employed to efficiently segment the areca nut.

The classification approach included six geometric parameters (principal axis length, secondary axis length, axis number, area, perimeter, and compactness of the areca nut image), three color features (mean grey level of an areca nut image on the R, G, and B bands), and defects area. To sort the quality of areca nuts, a back-propagation neural network classifier was used. The methods described here is effective for categorizing areca nuts with a 95.9 percent accuracy.

***Keyword:*** areca nut, classification, segmentation, detection, masking, mean, variance, data sets

## I. INTRODUCTION

For some Taiwanese, the areca nut is a popular and significant crop. In Taiwan, the annual output value of areca nut exceeds NTD 100,000 million. Despite the fact that chewing areca nut causes mouth cancer [1], the areca nut is commonly referred to as Taiwanese gum in Taiwan. Areca is commonly grown on steep slopes with good air circulation. Pathogens such as fungus, bacteria, viruses, and dangerous insects frequently infect Areca nuts. If the areca nut's surface is damaged, the price will drop. In Taiwan, traditional labor has been used to sort it thus far. Farmers' income is always impacted by the expense of human labor and sorting time.

Image processing is a strong technology that is commonly used in agricultural product detection. Color, geometric, and texture aspects are frequently employed to assess image quality. Park et al. [2] suggested a content-based picture classification technique based on texture parameters such contrast, diagonal moment, energy, entropy, homogeneity, second diagonal moment, and uniformity as input nodes to a neural network. According to nine textural parameters, Hsieh et al. [3] employed a neural network to detect the development stage of head cabbage seedlings. To identify flaws in cherries, Guyer and Yang [4] used genetic artificial neural networks and spectrum imaging. Indeed, neural networks, color, and texture variables were often used in plant and crop categorization [5–12]. Huang and Lin [13] developed methods to evaluate the geometric properties of Phalaenopsis seedlings using the boundary chain-code, Hoteling transformation, golden section search technique, and Bayesian classification. In the sorting procedure, such geometric features were applied. To segment banana leaves affected with illnesses, Camargo and Smith [14] employed color modification, image enhancement, and a determined optimal threshold. To extract disease spot regions and determine leaf areas, Shen et al. [15] used the Otsu approach, the HSI color system, and the Sobel operator.

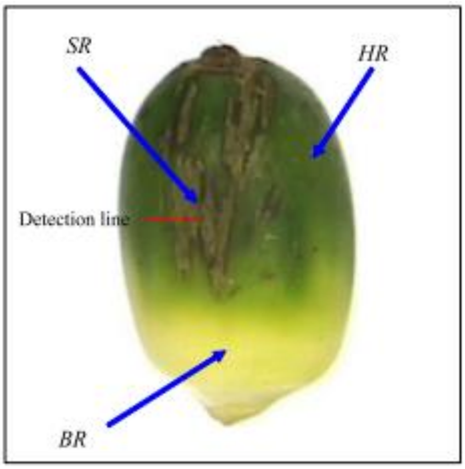


***Fig No.1: A Sample Picture of Areca Net***

In Taiwan, the labor cost for inspecting blemished areca nuts has climbed to up to 70% of the entire cost. As a result, farmers face high costs and should use a newly created automatic detecting and classifying technology to optimize their sorting process as soon as feasible. As a result, the goal of this research is to create a machine vision system that can detect and categories diseased or insect-infested areca nuts. The technological aims are to create an algorithm to extract color characteristics, geometric features, and fault areas from areca nuts, and then use those features to identify different classes.

## II. LITERATURE SURVEY

Based on our review of the literature, we discovered that classification of fruits, flowers, seeds, and other plant parts has been done; however, to our knowledge, no work has been done to classify areca nuts using an image processing technique. Four varieties of areca nuts are explored in this research. Api, Bette, Mine, and Gorublu are the names of the characters. The nuts are divided into just two classes for categorization purposes: api,bette, and mine belong to one class called Boiling Nuts (BN), while gorublu belongs to another called Non Boiling Nuts (NBN). Areca nuts may be categorized using our method based on their color component.

We did a study of 20 places and came up with the following findings:

- The hue of all gorublu arecanuts will be reddish yellow (Belongs to NBN class).
- Bette, api, and mine are all green in hue. (It belongs to the BN class)
- Nuts that are transitioning from bette to gorublu and have 25% of the green hue are likewise classified as BN.

***PROPOSED METHODOLOGY***

Areca nuts are categorized in this study based on their color components. In YCBCR color space, the RGB picture is transformed. The red and blue color components of areca nut objects were manually cropped from the photos, and the UCL (upper control limit) and LCL (lowest control limit) of the colors were calculated empirically using equations (1), (2), (3), and (4).

$$UCL_{cb} = \mu_{cb} + 3\sigma_{cb} \qquad (1)$$
$$LCL_{cb} = \mu_{cb} + 3\sigma_{cb} \qquad (2)$$
$$UCL_{cr} = \mu_{cr} + 3\sigma_{cr} \qquad (3)$$
$$LCL_{cr} = \mu_{cr} + 3\sigma_{cr} \qquad (4)$$

The mean intensities for blue (Cb) and red (Cr) chroma components are obtained using equations (5) and (6), respectively.

$$\mu_{cb} = \frac{1}{MXN}\sum_{i=1}^{M} \sum_{j=1}^{N} C_b(i,j) \qquad (5)$$
$$\mu_{cr} = \frac{1}{MXN}\sum_{i=1}^{M} \sum_{j=1}^{N} C_r(i,j) \qquad (6)$$

Where Cb & Cr are the color components of chromatic blue and red in the color space, and I & j are the pixel positions in rows M and columns N, respectively. The standard deviations for blue and red chroma components are Cb and Cr, respectively. Equations (7) and (8) are used to calculate these figures (8).

$$\sigma_{cb} = \sqrt{\frac{1}{MXN}\sum_{i=1}^{M} \sum_{j=1}^{N} C_b((i,j) - \mu_{cb})^2}$$
(7)

$$\sigma_{cr} = \sqrt{\frac{1}{MXN}\sum_{i=1}^{M} \sum_{j=1}^{N} C_r((i,j) - \mu_{cr})^2}$$
(8)

Before segmenting the item in the picture, the shadow was eliminated. Three control limits [11] are used to set the background colour to black for this assignment, and then the congested background is cleared using equation (8).

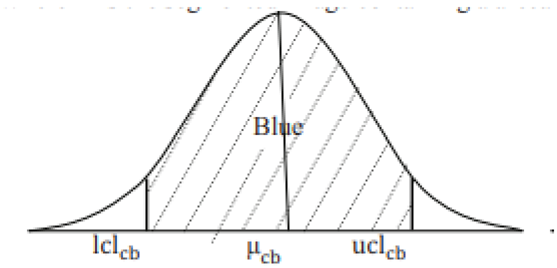


***Fig No.2: Blue chroma component control limits***

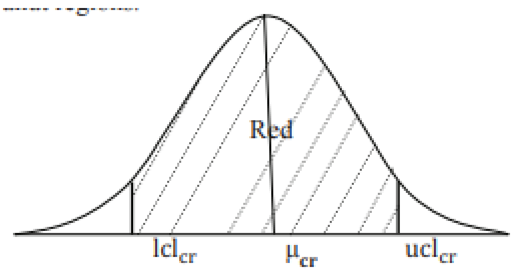


***Fig No.3: Red chroma component control limits***

Using equation, the segmented picture is transformed to a binary image and the morphological opening operation [12] is applied to eliminate any noise outside the areca nut object (9).

$$A\,B = (A \ominus B) \oplus B \qquad (9)$$

The segmented binary picture is A, and the structuring element is B. The 4X4 symmetric square structural element is explored in our experiment. Equations (10) & (11) are used to fill the various holes inside the areca nut object using a hole filling method [12].

$$X_k = (X_{k-1} \oplus B) \cap A_c \quad K = 1,2,3\ldots$$
$$m = X_{k\cap A}$$

Xp is a placeholder for Xp. All of the filled holes are contained in the set Xk. All of the filled holes and their bounds are contained in the set union Xk and A. If Xk=Xk-1, the algorithm is finished. A symmetric structural element is designated by the letter B. After filling the gaps in the image mask, the areca nut object will have the same form and size as the original picture in figure above.

## III. MATERIALS AND METHODS

***Image Acquisition System***

Images of areca nuts were captured using a machine vision system. A GigE CCD (coupled-charge device) color camera with a zoom lens (DFK-31AG03, Imaging Source Inc.) and a personal computer (Intel Pentium 4 CPU 2.4 GHz) are included in this setup. To get RGB color pictures of 640 480 pixels, the Open Source Computer Vision Library (OpenCV 1.0, Intel Corporation) was connected to the applications. The picture acquisition was done with a CCD camera set to 4600 lx and an aperture of F4.0 (iris diaphragm). Images were saved in the tagged image file (TIF) format on a computer's hard disc. Microsoft Visual C++ 6.0 was used to process the images.

***Features Extraction***

Following the extraction of the areca nut's characteristics, segmentation of the full picture of the areca nut is required. The thresholding, hole-filling, closing, and opening processes segment the complete picture of an areca nut [18]. To begin, the areca nut's primary axis must be recognized. Assume that the binary image of an areca nut is f(xi, yi) (where I = 1, 2,..., m, and the total number of pixels is m), and that the total number of pixels is m. The centroid can be calculated using the formula

$$X = \sum_{i=1}^{m} \frac{x_i}{m} \text{ and } Y = \sum_{j=1}^{n} \frac{\mathrm{Yj}}{n}$$

The covariance matrix can be calculated using the formula

$C = \sum_{i=1}^{m} U_i U^t/m - MM^t$ , where Ui is the ith coordinate of the vector

***Spot Region Detection***

In the appearance of the areca nut, there are healthy regions (HR), base regions (BR), and spot regions (SR), as illustrated in Fig. 1. Once the injured areca nut has been identified, efficiently segmenting SR is critical. The difference between SR and HR, on the other hand, is difficult to identify using the thresholding approach. The grey level histogram, for example, cannot be used to determine whether or not to extract SR using the thresholding decision rule. The threshold values (T) 160, 150, and 80 are used to segment SR on the red (R), green (G), and blue (B) bands, respectively (Fig. 2(a)–(c)). Figure 2(g) depicts the segmentation findings (i). As a result, a unique approach for SR segmentation of areca nut in the study needs to be developed.

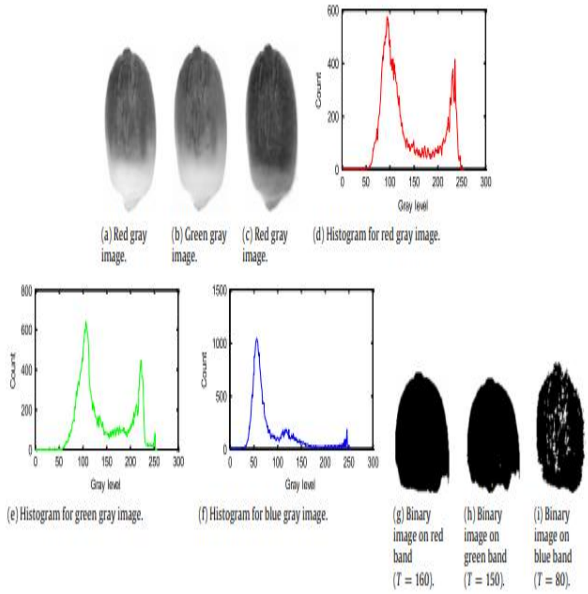


***Fig No.4: The steps involved in the thresholding technique for areca nut photos. (The reader is directed to the web version of this article for interpretation of the colour references in this figure legend.)***

This experiment was preceded by the following. To begin, a detection line with a scanning resolution of one pixel was used to discover grey levels in an image of an areca nut. As shown in Fig.

*ISSN: 2248-9622, Vol. 12, Issue 2, (Series-II) February 2022, pp. 53-59*

3, the grey level distribution shapes in HR, SR, and BR are all distinct. On the HR area, the grey level curves on R and G bands are relatively near, as illustrated in Fig. 3. (a). There were even some R, G, and B curves that overlapped (the sites C1, C2, and C3). On SR, the distinct grey level zones are clearly visible. In addition, as shown in Fig. 3, there are several grey levels that developed on BR (b).

Second, in order to resolve the overlapping circumstances, the areca nut picture was handled with the smoothing operator [18] (as shown in Fig. 4). A detection line may be used to provide grey level distributions on the SR, HR, and BR, as shown in Fig. 5. There are notable variances, as illustrated in Fig. 5, where the sites C1, C2, and C3 on the SR and HR were achieved following smooth operation. As a result, the grey levels fluctuation on the R, G, and B bands may be used to distinguish SR, HR, and BR.

As a result, a novel application—the detection line (DL) algorithm—was successfully implemented to extract SR, HR, and BR for the smoothing picture based on the grey level differences on the R, G, and B bands. The following is a description of the DL algorithm:

1. Smooth operator transforms the original areca nut picture p(x, y) into an image f(x, y).
2. Calculate the grey level distribution on the R, G, and B bands using the detection line (with scanning resolution in pixels) that travels along the areca nut's principal axis.
3. On a steady state basis, distinguish between SR, HR, and BR (i.e., light, iris-diaphragm).

## IV. CLASSIFICATION

In the categorization procedure, geometric and color characteristics analyses have been frequently used. Geometric and color criteria were used to classify the quality of areca nuts in this study. Six geometric parameters (principal axis length, secondary axis, axis number, area, perimeter, compactness), three cooler features (mean grey level on R, G, and B bands: Rmean, Gmean, and Bmean), and the SR area were used to classify the quality of areca nut.

When a pattern recognition method is required, the artificial neural network (ANN) has several uses. A back propagation neural network (BPNN, as illustrated in Fig. 6) was used to categories areca nuts into excellent, good, and terrible categories in this study. There are three layers in the BPNN classifier: an input layer, a hidden layer, and an output layer. The input layer contains 10 nodes relating to the SR region, 3 color features, and the aforementioned 6 geometric characteristics. The characteristics in the input are normalized between 0 and 1. The output layer consists of nodes categorized as Excellent (E), Good (G), and Bad (B) (B). The number of nodes nh in the hidden layer was first estimated using the algorithm below.

$$n_h = [(n_i + n_0)/2] + (n_p)^{0.5}$$

The number of input nodes is ni, the number of output nodes is no, and the number of input patterns in the training set is np. The goal of the learning process is to discover a relationship in a pattern created by the characteristics of each individual.

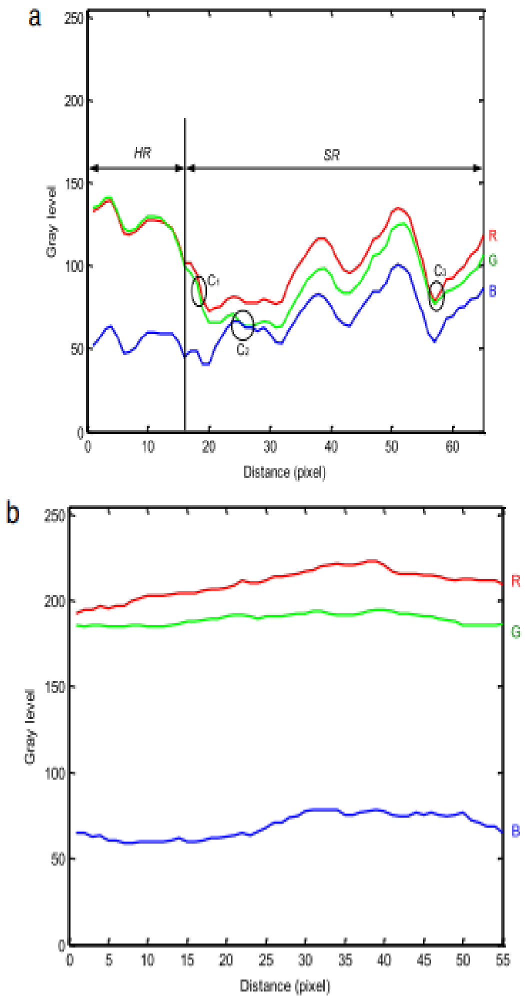


***Fig No.5: The grey level distribution for an original areca nut picture. HR and SR, to begin with. BR (b)***

the areca nut The weights of the BPNN are modified until the error convergence criteria is set at 0.1. The indication for an error is delivered by

$$E_t = \sum_{p=1}^{3} e_p(t) = \sum_{p=1}^{3} d_p(t) - y_p(t)$$

Image processing methods, the DL algorithm, and the BPNN classifier were used to create an algorithm for detecting and classifying areca nuts. The following is a description of the algorithm:

Step 1: Estimate the areca nut's three hues and six geometric characteristics.
Step 2: SR area estimate and segmentation
Step 3: To classify the quality of areca nuts, create and test a BPNN classifier.

Figure 5 depicts the stages involved in defect region identification and areca nut categorization. Microsoft Visual C++ 6.0 and OpenCV 1.0 were used to create the software that detected and classified the areca nuts.

## V. RESULT

For this study, the most common commercial variety of arecanut is used. There are 629 photos in the database from 15 distinct places, with 71 being used for training and the rest 558 for testing. A Canon digital colour camera (Power Shot A1100IS) was used to capture images at a resolution of 3000x4000 pixels. In natural daylight, all of the images were shot to roughly fill the camera's field of view with white back ground.

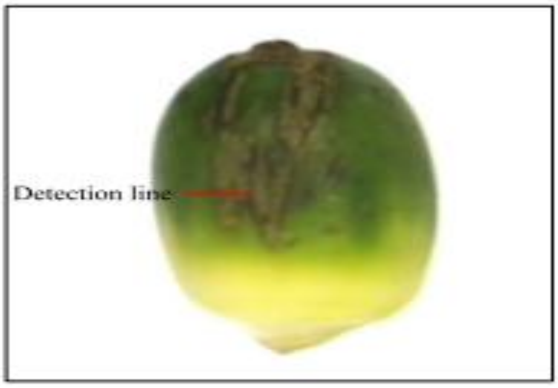


***Fig No.6 : After the smoothing procedure, the picture.***

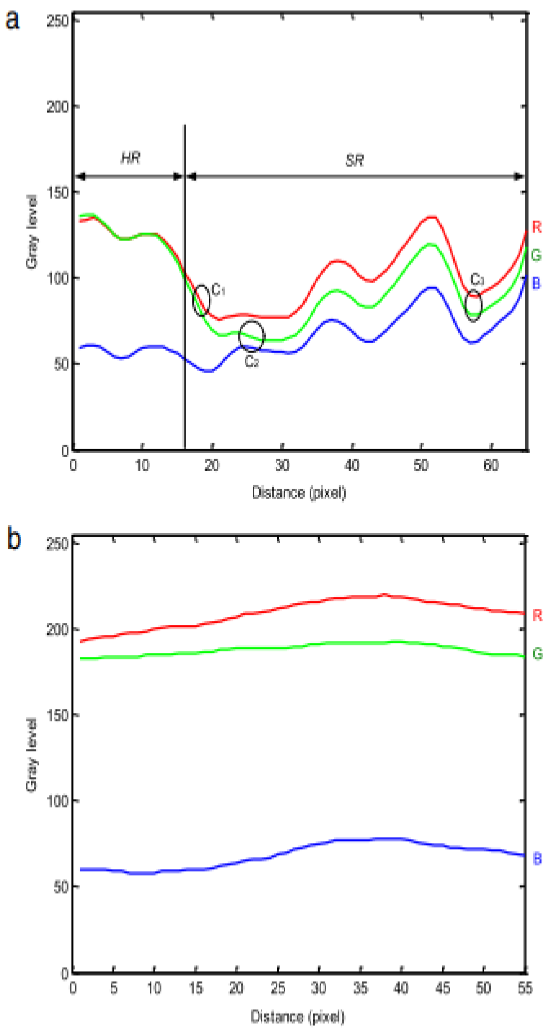


***Fig No.7 : After smoothing, the distribution of grey levels for the areca nut.***

The CCD camera, for example, obtained a picture of an areca nut, as seen in Fig. 9. (a). The R, B, and G color bands were then isolated from the original picture (Fig. 9(b)–(d)). The binary picture on the B band (seen in Fig. 9(e)) was created using a 180 is the cutoff value. For the binary and smooth pictures (Figs. 9(e) and (f), a segmented image utilizing an AND logic operator is shown in Fig. 9(g). As shown in Fig. 9(g), an isolated SR picture was retrieved using the DL method and image processing techniques (such as hole-filling, erosion, dilation, opening, and closing operators) (h). According to the above-mentioned verification, the suggested DL algorithm can estimate the SR area of an areca nut.

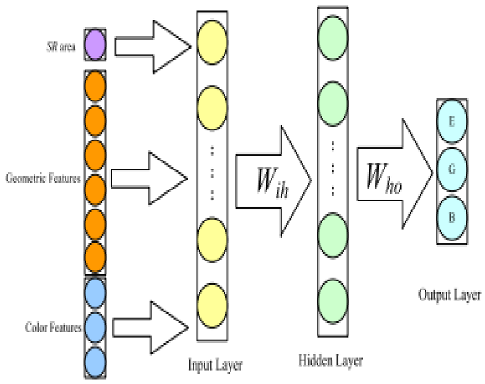


**Fig. 6.** The structure of the BPNN classifier.

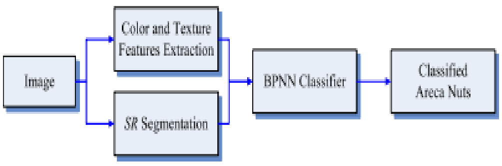


***Fig No.8: Detection and Classification Stages***

The BPNN classifier was built using 144 samples, including 49 Excellent, 46 Good, and 49 Bad, which were randomly picked from 287 photos using the 50 percent –50 percent splitting method. Using ten input features, three output categories, and 144 input samples, Eq. (1) yielded eighteen hidden nodes. Matlab 8.0 routines were used to construct the BPNN classifier. The set of training samples for each configuration was chosen at random until the BPNN converged. When the training set contains some inaccurate samples in the BPNN, over-fitting is common.

Over fitting was unlikely to occur since the grades of areca nuts in training samples were previously known prior to the training procedure. The maximum number of iterations was set at 10,000 to guarantee that the impact of overfitting was minimal throughout the training phase. Further research revealed that when the ratio of error convergence criterion was less than 0.1, the same categorization was attained. To test the system, 48 Excellent, 412 Good, and 178 Bad photos were randomly selected. The classification accuracies for Excellent, Good, and Bad grades, respectively, are 98.37 percent, 89.1 percent, and 96.72 percent, and the findings are shown in Table 1. The accuracy rate was 98 percent on average. There were 130 valid classifications and 13 incorrect categories in all.

For faster processing, images were downsized to a resolution of 150x200 pixels. As shown in table 1, the proposed technique efficiently classifies various areca nuts such as api, bette, mine, and gorublu into two classes: BN and NBN. The dirty and bruised nuts are to blame for the misclassification.

| Class | Testing Samples | misclassification | Success Rate |
|---|---|---|---|
| BN | 412 | 5 | 98.36 |
| NBN | 178 | 5 | 96.72 |

The suggested method can reliably and effectively identify spots and categories the quality of areca nuts using a CCD camera in this investigation, however it is unable to examine for covered blades (or other face). We will develop a practical design for an autonomous system to classify areca nuts in the future. An auto-feeding mechanism, a reversed orientation device (for example, a rotation mechanism), an auto-discharge mechanism, and detecting software make up the system.

## VI. CONCLUSION

We utilized control limits to segment the areca nuts from the picture in this article. The blue color component is generally muted in the segmented region, and only the red and green color components are utilized to categories the areca nuts. Three sigma variances from the mean are used to calculate the upper and lower control limits of the blue and red chroma color components. The areca nut areas can be segmented using these control limitations. The red and blue color components of the segmented section of the areca nut are used to further classify it. The proposed method's efficacy was demonstrated by experimental data. This approach may be used to other things, such as fruit, seed, and flower categorization, where sorting and quality grading is usually done by professionals.

## REFERENCE

[1]. B.J. Dave, A.H. Trivedi, S.G. Adhvaryu, Role of areca nut consumption in the cause of oral cancers, a cytogenetic assessment, Cancer 70 (5) (1992)1017–1023.
[2]. C.L. Hsieh, S.F. Cheng, T.T. Lin, Application of image texture analysis and neural network on the growth stage recognition for head cabbage seedlings, Journal of Agricultural Machinery 6 (2) (1997) 1–13 (in Chinese).
[3]. D. Guyer, X. Yang, Use of genetic artificial neural networks and spectral imaging for defect detection on cherries, Computers and Electronics in Agriculture 29 (2000) 179–194.
[4]. J.A. Marchant, C.M. Onyango, Comparison of a Bayesian classifier with a multilayer feed-forward neural network using the example of plant/weed/soil discrimination, Computers and Electronics in Agriculture 39 (2003) 3–22

[5]. D. Morquin, M.B. Ghalia, S. Bose, An integrated neural network-based vision system for automated separation of clods from agricultural produce, Engineering Applications of Artificial Intelligence 16 (2003) 45–55.
[6]. D. Moshou, C. Bravo, J. West, S. Wahlen, A. McCartney, H. Ramon, Automatic detection of ‘‘yellow rust’’ in wheat using reflectance measurements and neural networks, Computers and Electronics in Agriculture 44 (2004) 173–188.
[7]. K.Y. Huang, Application of artificial neural network for Phalaenopsis seedling diseases using color and texture features, Computers and Electronics in Agriculture 57 (1) (2007) 3–11.
[8]. K.Y. Huang, T.C. Lin, Estimating the geometric characteristics of Phalaenopsis orchid during big plant stage with machine vision, Journal of Agricultural Machinery 9 (2) (2000) 13–26 (in Chinese).
[9]. S.L. Chen, S.T. Chou, Mura defect detection using discrete wavelet transform, Journal of the Chinese Society of Mechanical Engineering 30 (1) (2009) 75–80.
[10]. J.H. Park, G.S. Lee, S.Y. Park, Color image segmentation using adaptive mean shift and statistical model-based methods, Computers & Mathematics with Applications 57 (6) (2009) 970–980.
[11]. Majid Rashidi And Keyvan Seyfi. Classification of Fruit Shape in Kiwifruit Applying the Analysis of Outer Dimensions. Int. Jr. of Agriculture & Biology; 2007, 1560–8530,09–5–759–762.
[12]. Michael Recce, John Taylor, Alessio Plebe & Giuseppe Tropiano. High Speed Vision-Based Quality Grading of Oranges. IEEE Xplore; 1996, 0-8186-7456-3.
[13]. P.S Hiremath and Ajit Danti, Detection of multiple faces in an image using skin color information and Lines-of-Separability face model. International Journal of Pattern recognition and Artificial Intelligence, World scientific Publisher; 2006,Vol. 20(1), pp. 39-61.
[14]. Salmet Riyadi, Mohd. Marzuki Mustafa, Aini Hussain and Azman Hamza, Papaya fruit grading based on size using image analysis. Proceedings of the International Conference on Electrical Engineering and Informatics Institute Teknologi Bandung, Indonesia; 2007, 17-19.
[15]. Son Lam Phung, Abdesselam Bouzerdoum and Douglas Chai. Skin Segmentation Using Color Pixel Classification: Analysis and Comparison. IEEE Transactions on Pattern Analysis and Machine Intelligence; 2005, vol. 27, pp. 148-154.
[16]. Zulham Effendi, Rizauddin R, Jaharah A Ghani, Zahira Yaakob. Development of Jatropha Curcas Color Grading System for Ripeness Evaluation. European Jr. of Scientific; 2009, 662-669.
[17]. Meftah Salem M Alfatni, Abdul Rashid Mohamed Shariff, Helmi Zulhaidi Mohd Shafri, Osama M Ben saaed and Omar M Eshanta. Oil Palm Fruit Bunch Grading System Using Red, Green and Blue Digital Number. Journal of Applied Sciences; 2008, 8(8):1444-1452.
[18]. Mohammad Saber Iraji and Azam Tosinia, Classification Tomatoes on Machine Vision with Fuzzy the Mamdani Inference, Adaptive Neuro Fuzzy Inference System Based (Anfis-Sugeno). Australian Journal of Basic and Applied Sciences; 2011, 5(11): 846-853.
[19]. Nimesh, S., Delwiche, M. J., Scott Johnson R., Image analysis methods for realtime color grading of stone fruit. Computers and Electronics in Agriculture 9(1); 1993, 71-84.
[20]. Laykin, S., Alchanatis, V., Fallik, E., Edany, Image processing algorithms for tomatoes classification. Trans. of the American Society of Agricultural Engineers 45 (3); 2002, 851-858